

# Fermi LAT Analysis: Likelihood Testing

Fermi Summer School 2012, Lewes DE  
Jeremy S. Perkins (FSSC)

# The FSSC Website

<http://fermi.gsfc.nasa.gov/ssc/>

The screenshot shows the homepage of the Fermi Science Support Center (FSSC). At the top left is the NASA logo. To its right, the text reads "National Aeronautics and Space Administration" and "Goddard Space Flight Center". On the top right, there is a search bar with "Fermi" entered and a "GO" button. Below the search bar, it says "FSSC • HEASARC • Sciences and Exploration". The main header features the text "Fermi Science Support Center" over a background image of the Fermi satellite in space. A navigation menu below the header includes links for "Home", "Observations", "Data", "Proposals", "Library", "HEASARC", "Help", and "Site Map". The main content area contains a paragraph describing the FSSC's role: "The Fermi Science Support Center (FSSC) ... the guest investigator program, creates and maintains the mission time line, provides analysis tools for the scientific community, and archives and serves the Fermi data. This web site is the portal to Fermi for all guest investigators." To the right of this text is a "Latest News" section with links to "Fermi Sky Blog" and "Fermi Blog", dated "May 1, 2012", and a headline "Cycle 5 Selected Programs Released". Below the news section is a "released" notice: "The FSSC is pleased to announce the availability of a new version of the Fermi Science Tools (v9r27p1) along with".

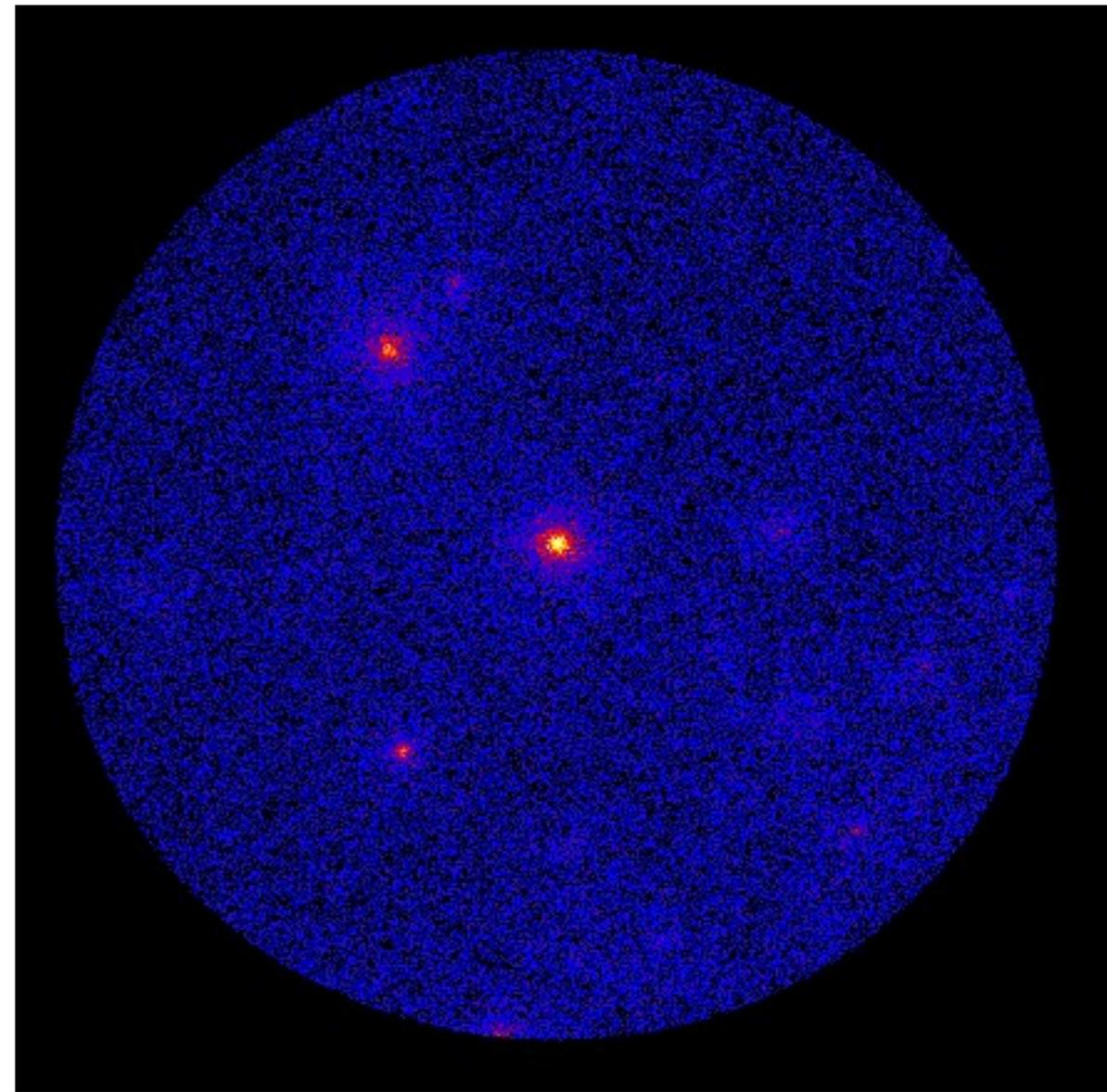
**Documentation, Software, Data, Threads**

**Contact Us, FAQ**

This all-sky view from Fermi reveals bright emission in the plane of the

# The Challenge

- Gamma-ray data is a list of counts (photons) reconstructed in the detector.
- Qualitative exploration of the data suggests the presence of sources (spatial clustering).
- Quantitative analysis requires evaluating the significance of a 'model' of our region.



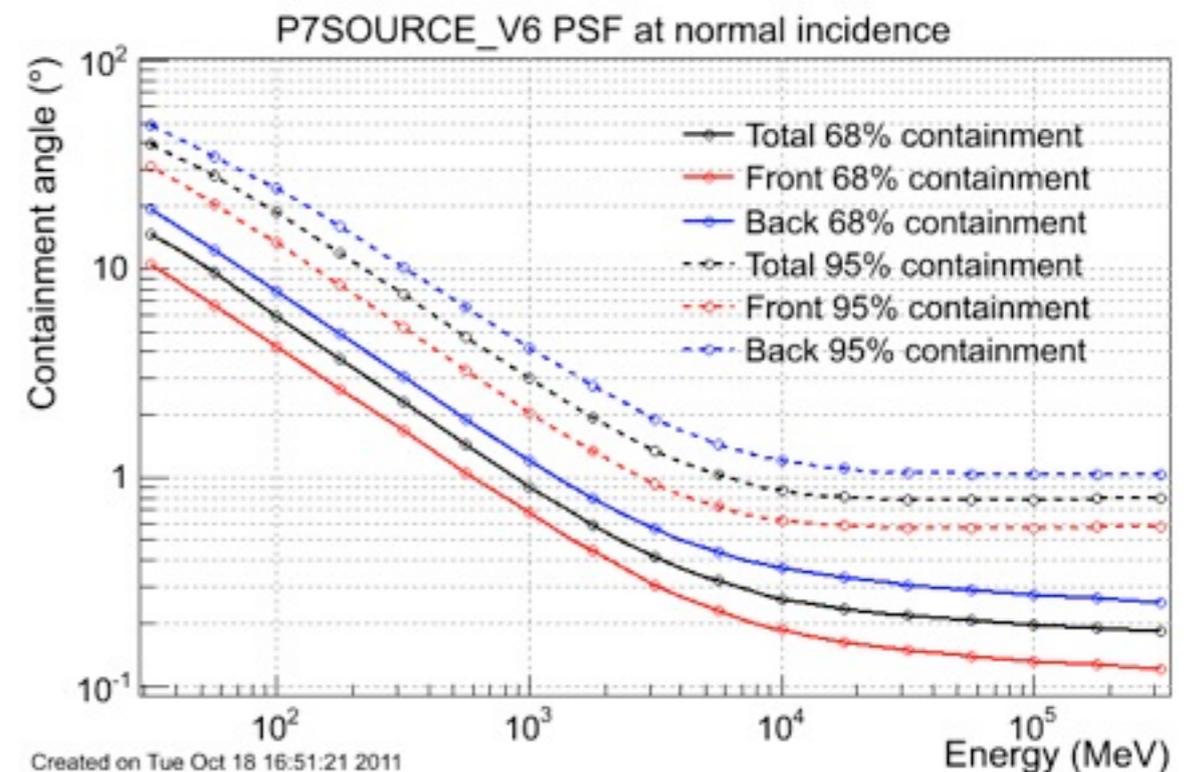
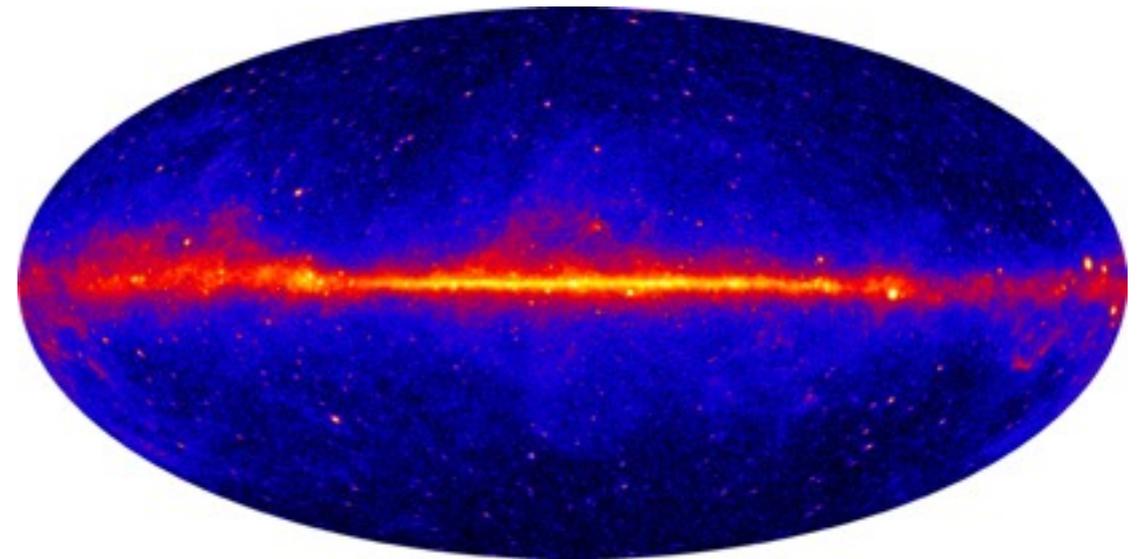
Count Map

# The Model

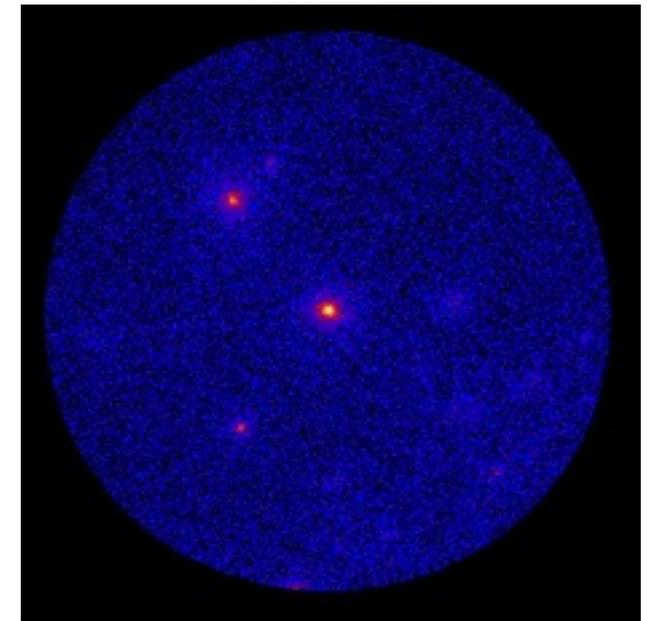
- Basically the initial ‘model’ is a guess of the various parameters of the sources in our region:
  - The location
  - The spectral shape
  - The flux
  - etc.
- The guess can be easier if you have a starting point like the 2FGL.
- We quantify (using the Likelihood Method) the statistical significance of the model and vary the parameters to determine the most likely parameter values.

# Why Model Fitting?

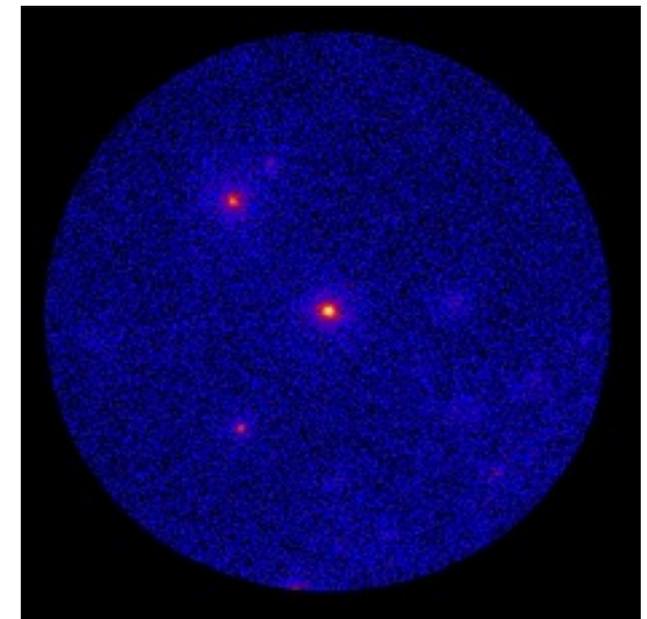
- We use the Likelihood method because the LAT data are limited by statistics, a bright diffuse background and a broad PSF.
- The model construction defines the questions we want to ask which means:
  - It will not answer a question you are not asking (ie. unknown parameters).
- The Likelihood will not tell you if a fit is 'good'. If the model does not represent your data well, the results will also not represent reality well.



# A Graphical Example

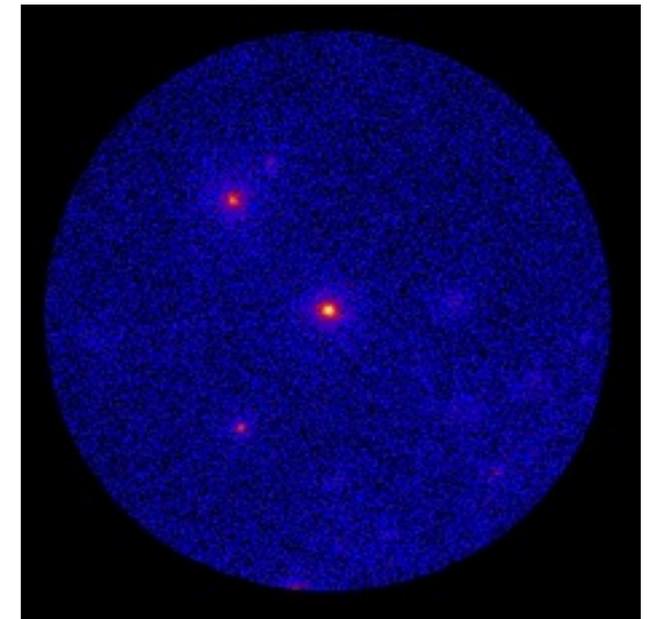
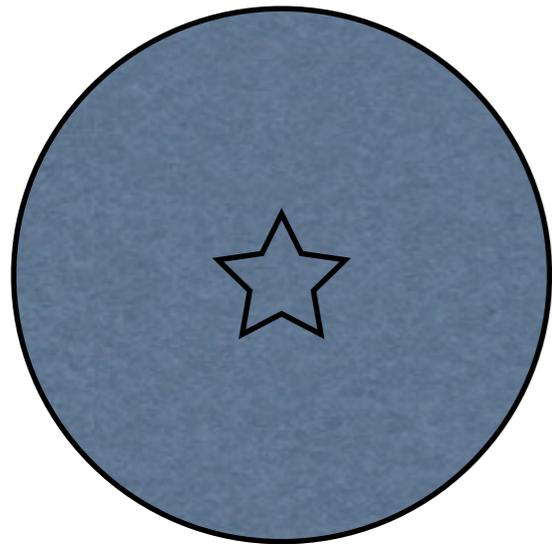


Our Data

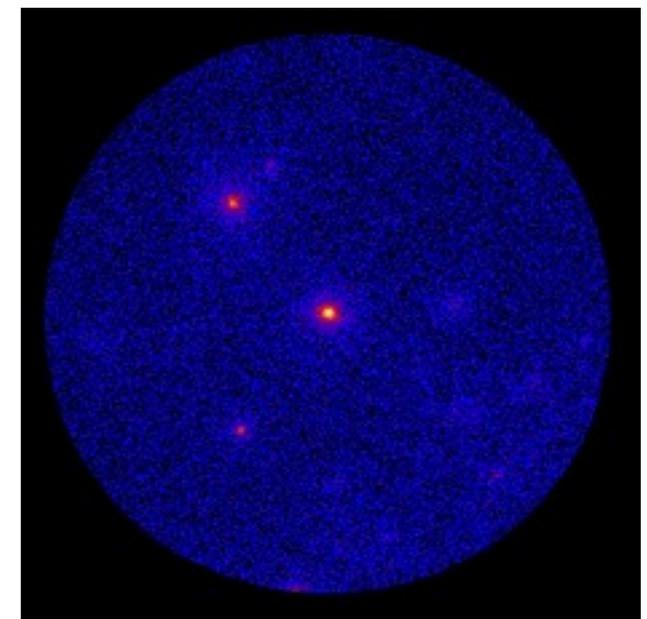


# A Graphical Example

Model 1

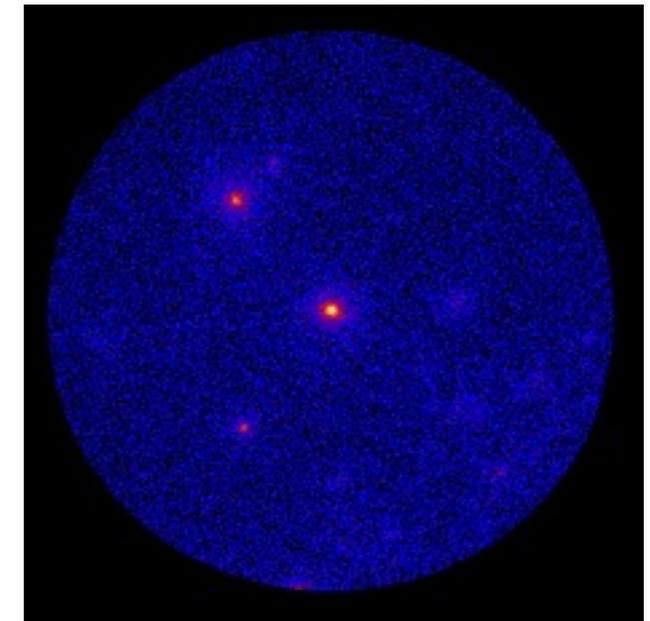
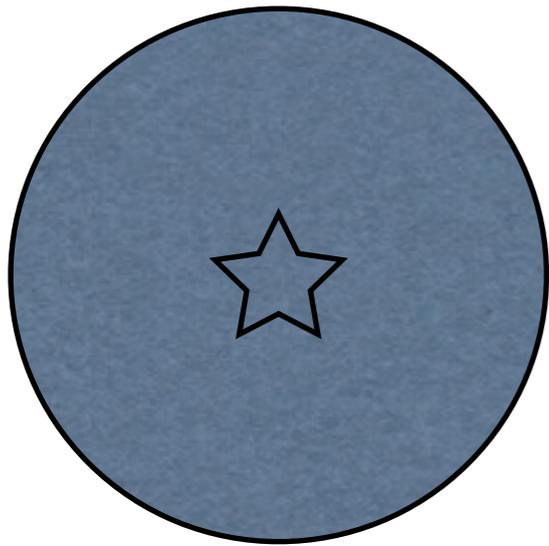


Our Data



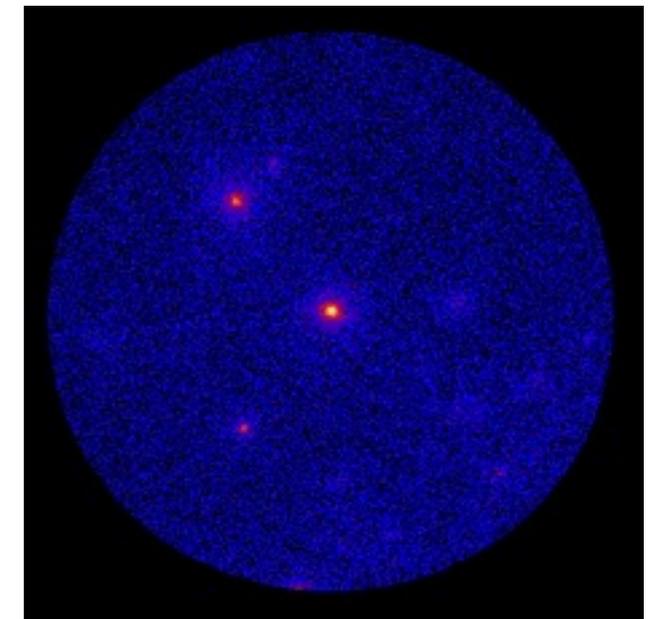
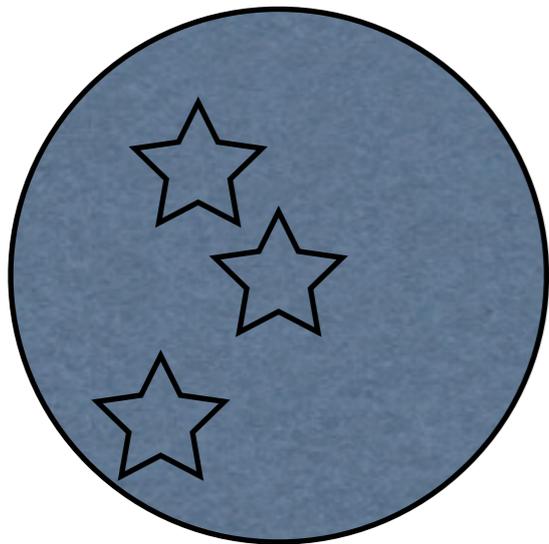
# A Graphical Example

Model 1



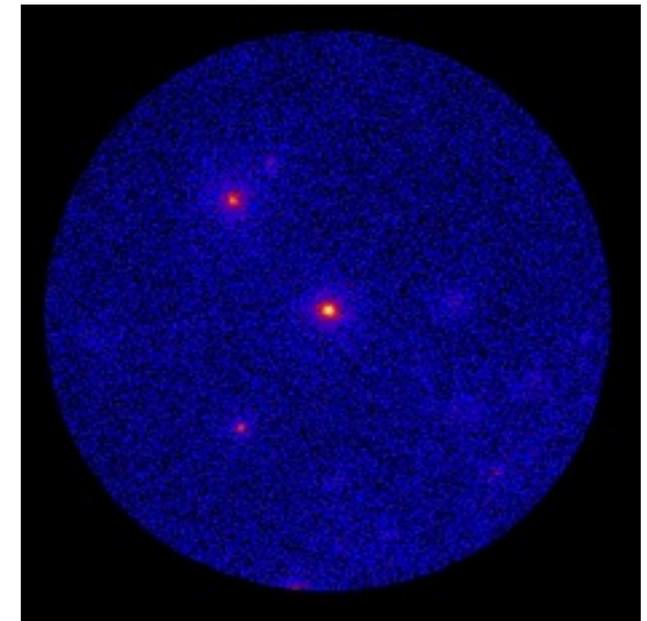
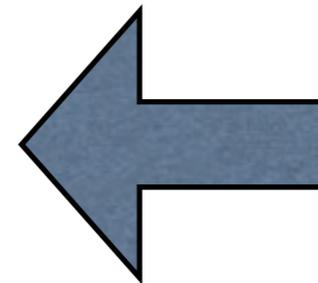
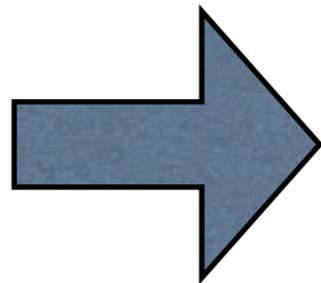
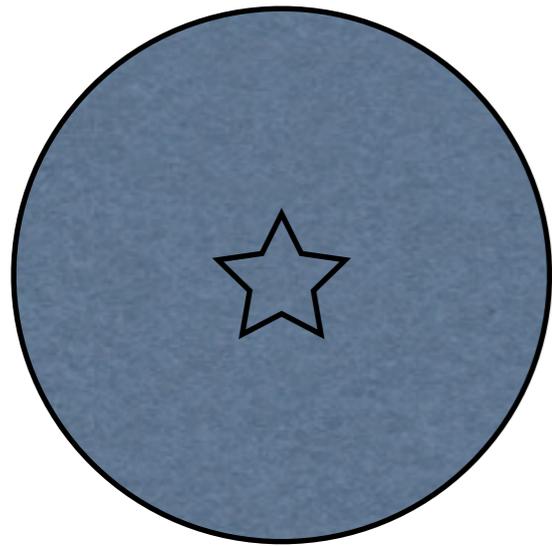
Our Data

Model 2



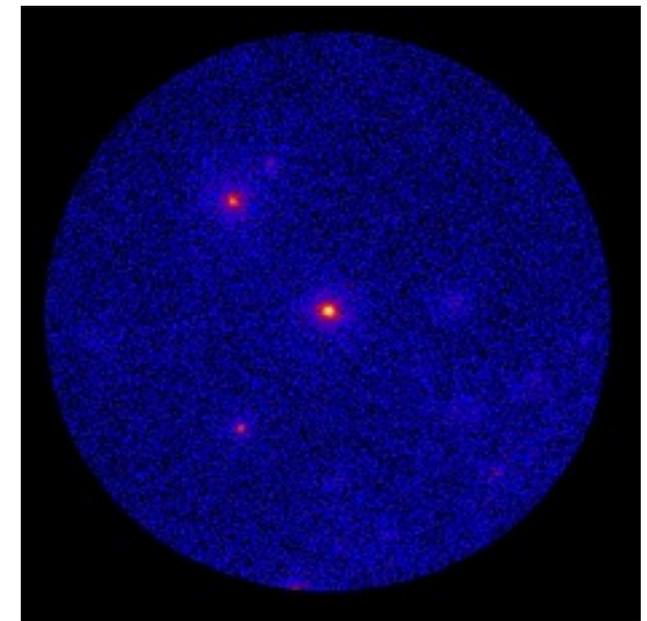
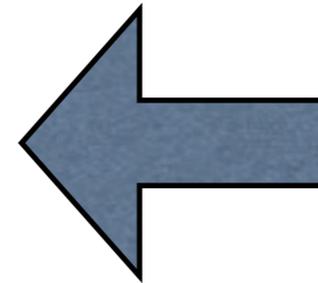
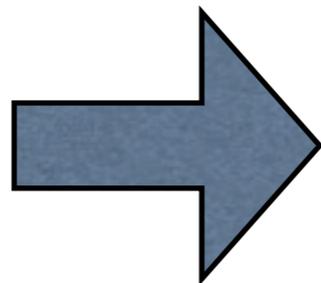
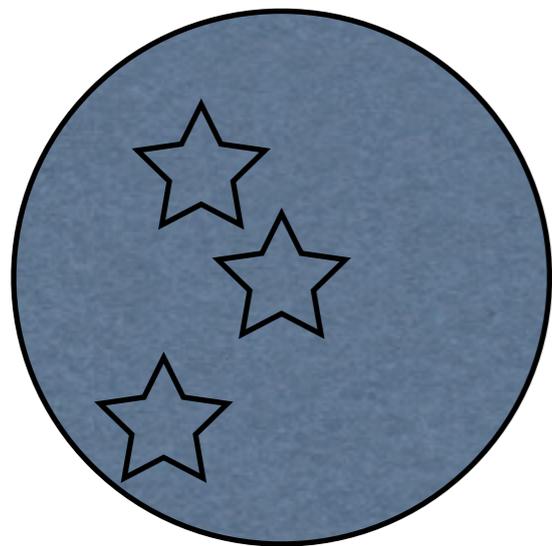
# A Graphical Example

Model 1



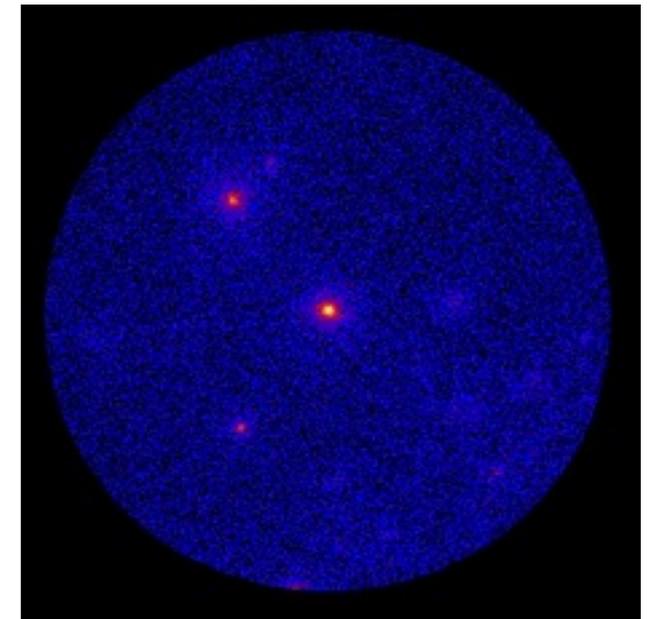
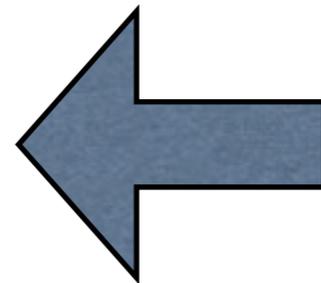
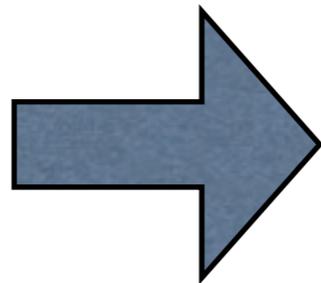
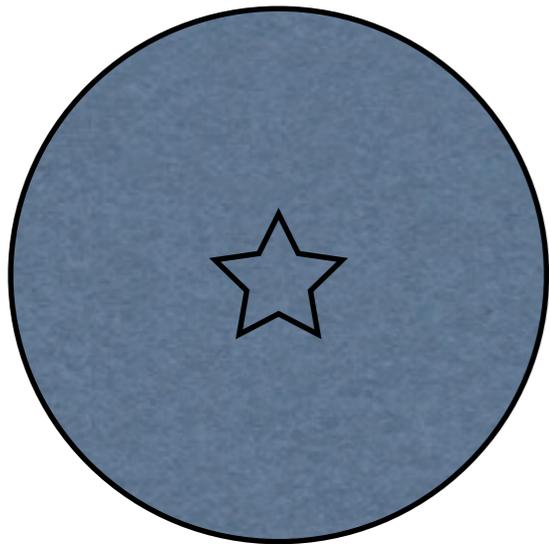
Our Data

Model 2



# A Graphical Example

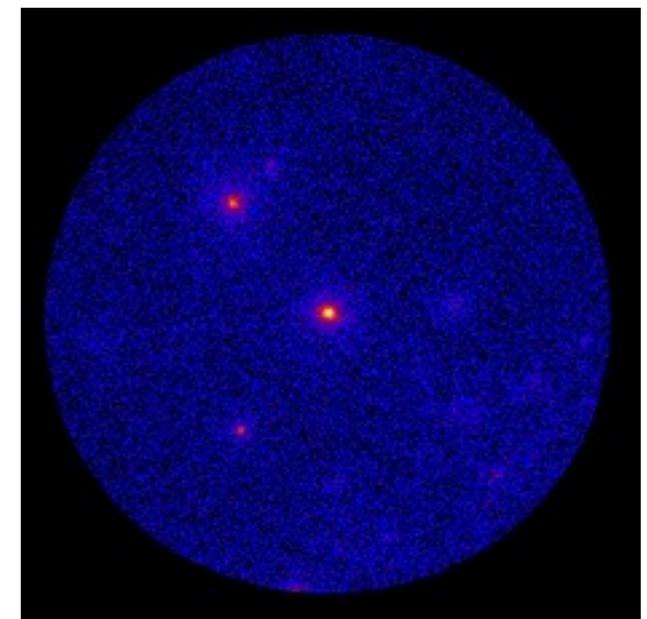
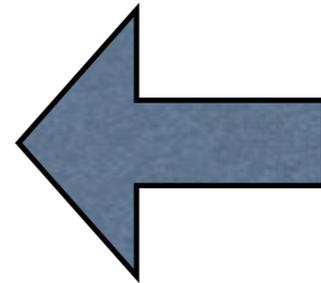
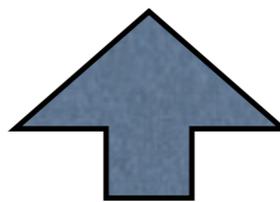
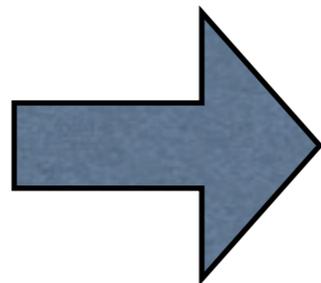
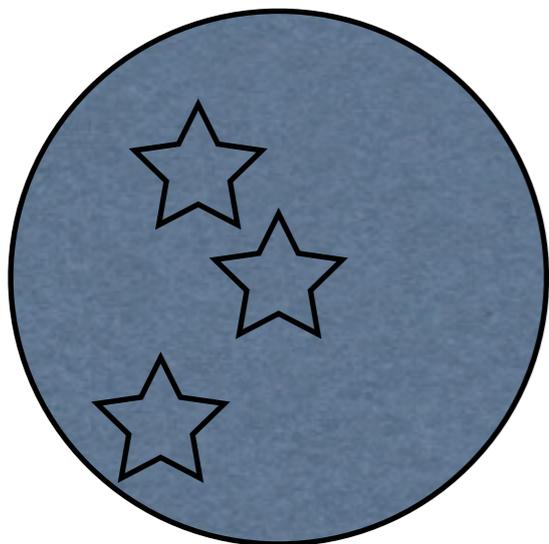
Model 1



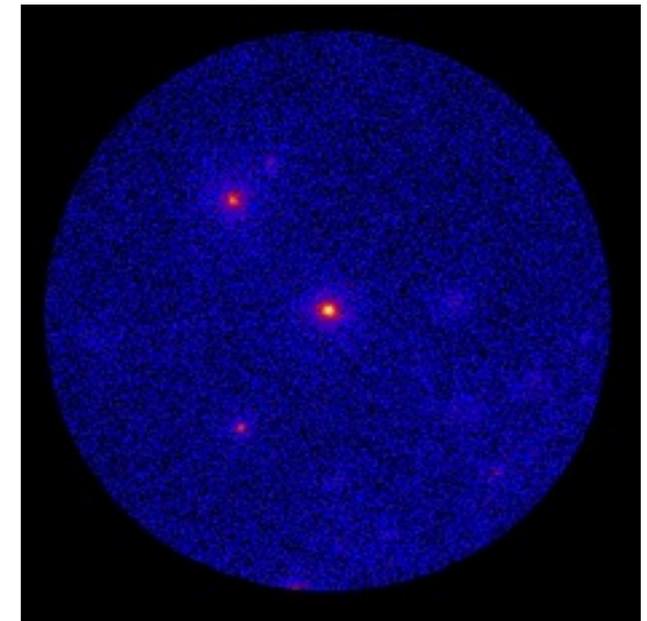
Model 2 > Model 1

Our Data

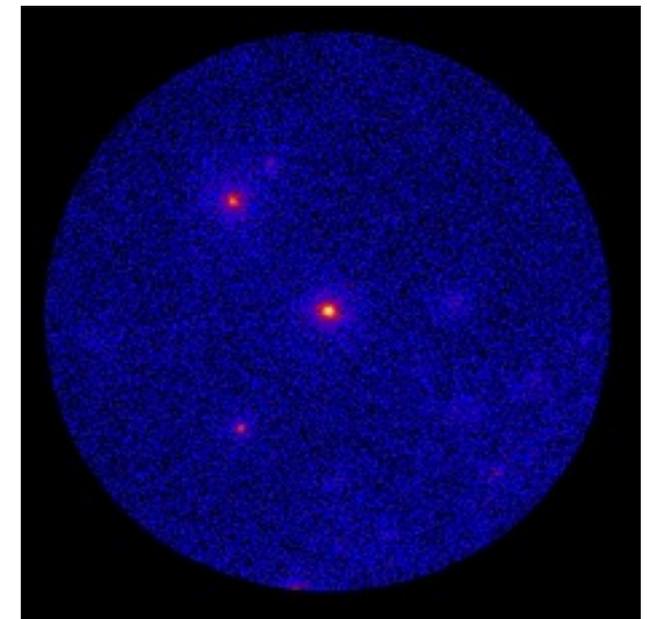
Model 2



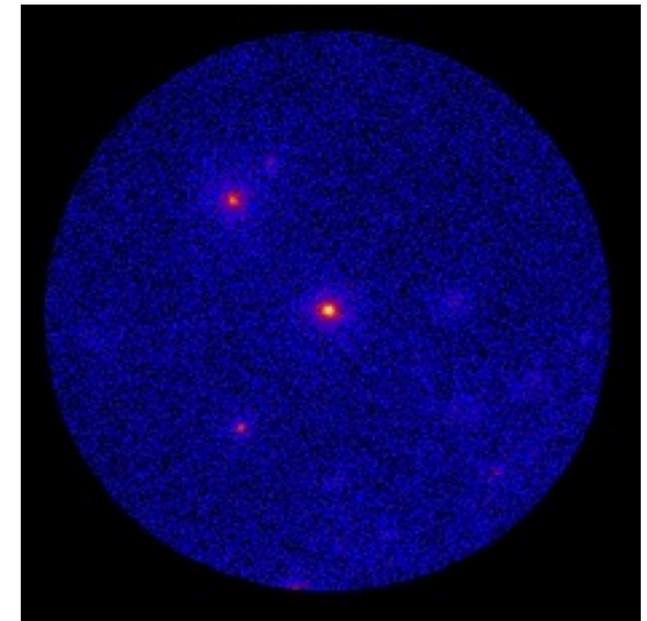
# A Graphical Example



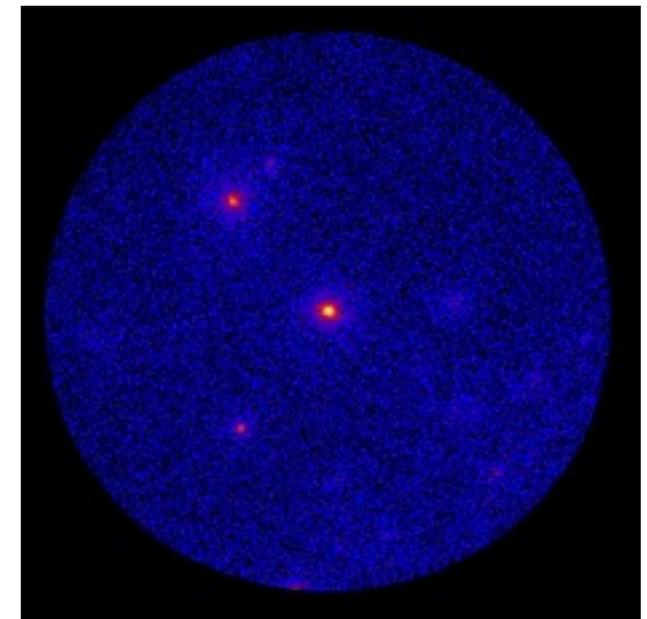
Our Data



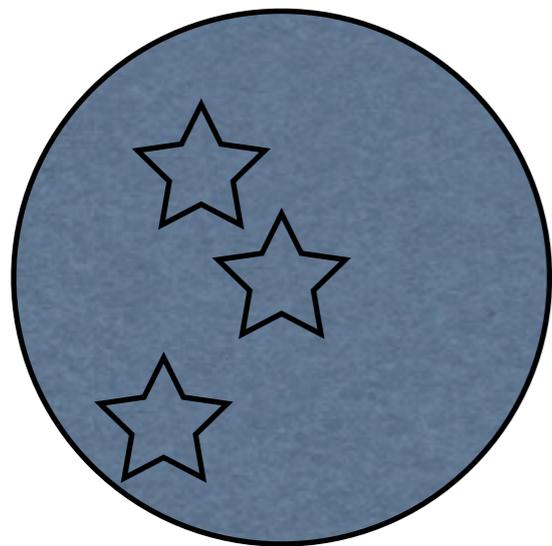
# A Graphical Example



Our Data

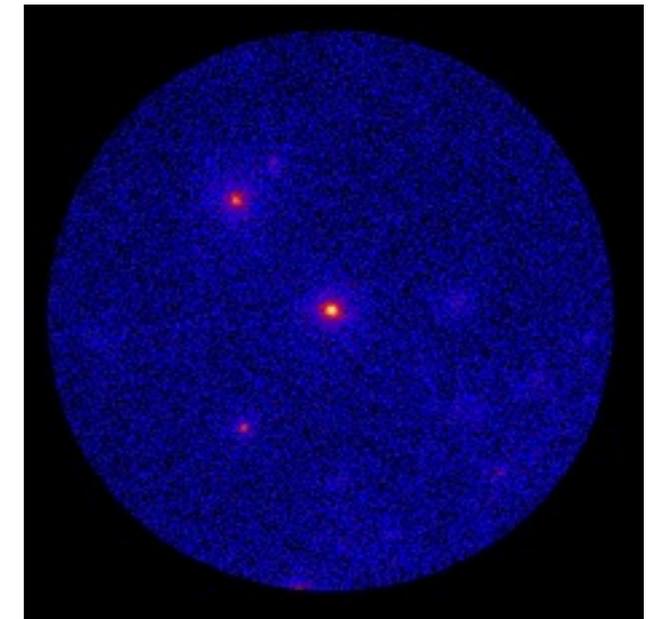
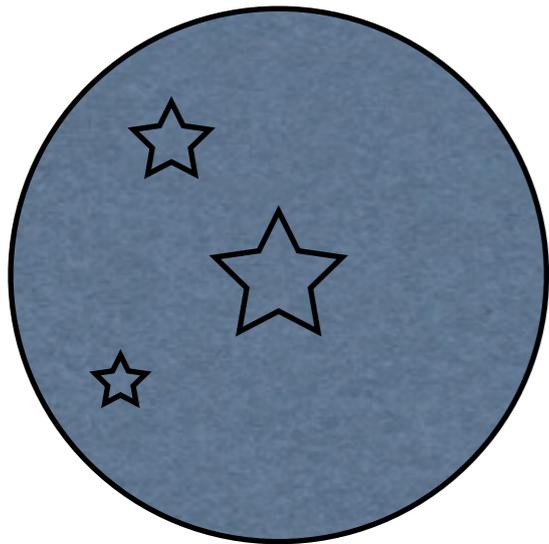


Model 2



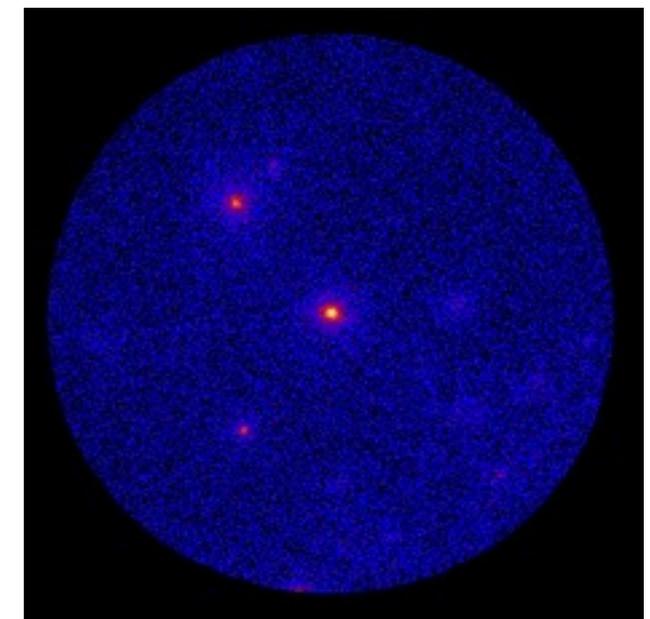
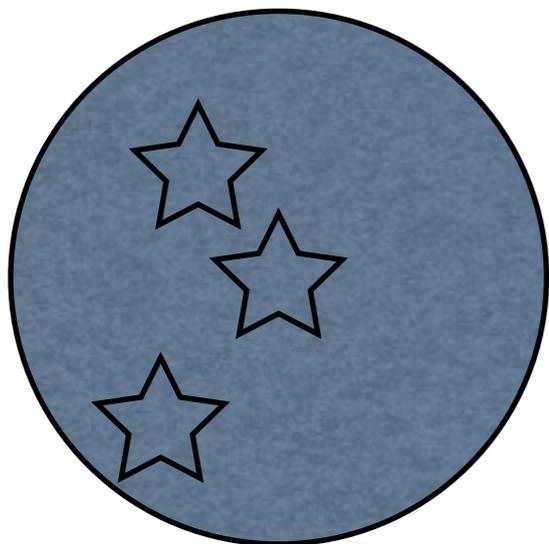
# A Graphical Example

Model 3



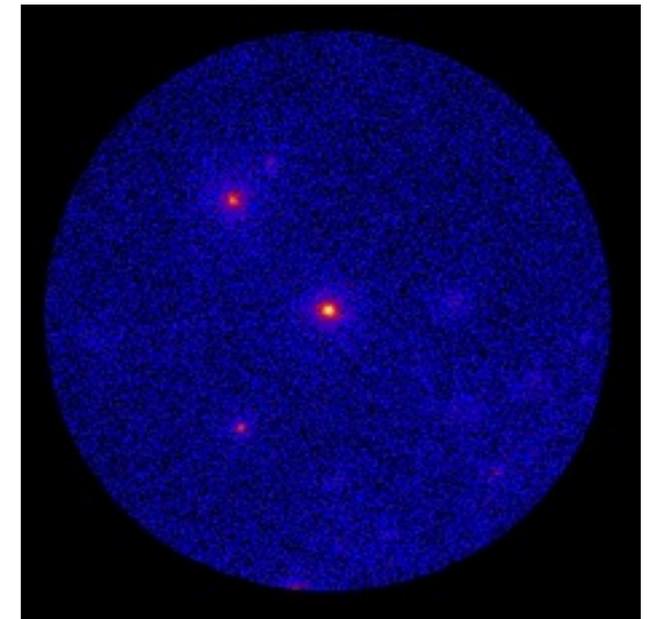
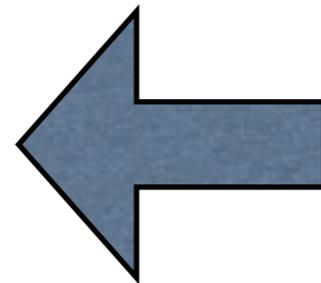
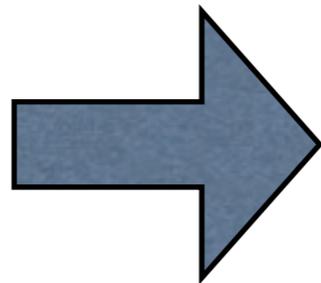
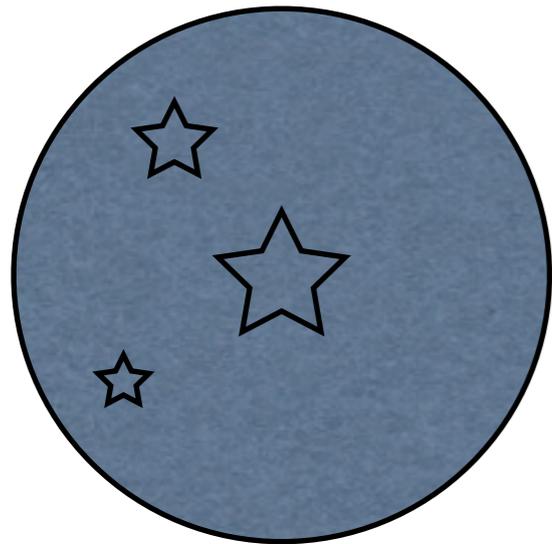
Our Data

Model 2



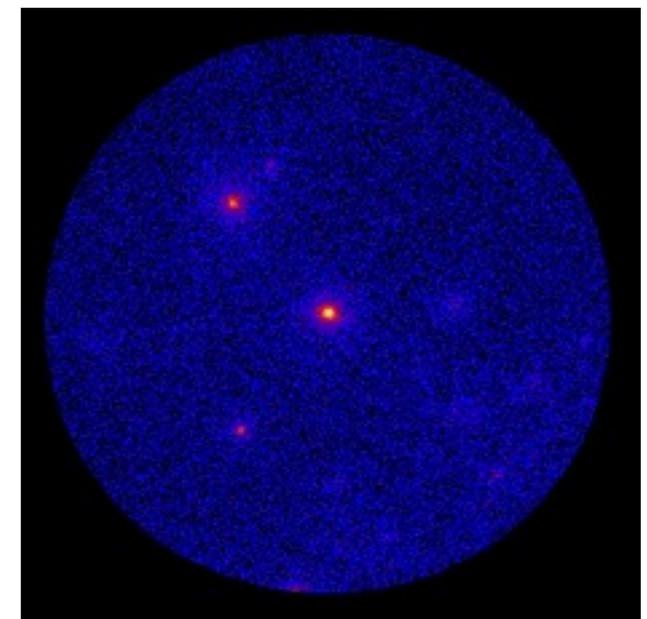
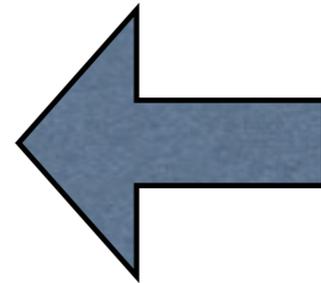
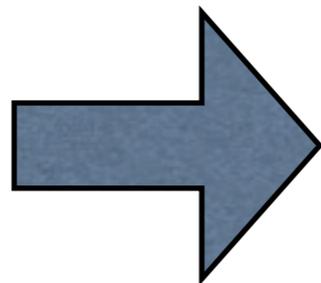
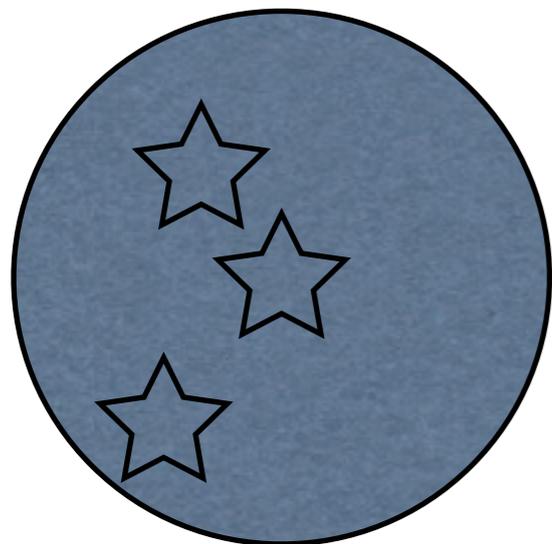
# A Graphical Example

Model 3



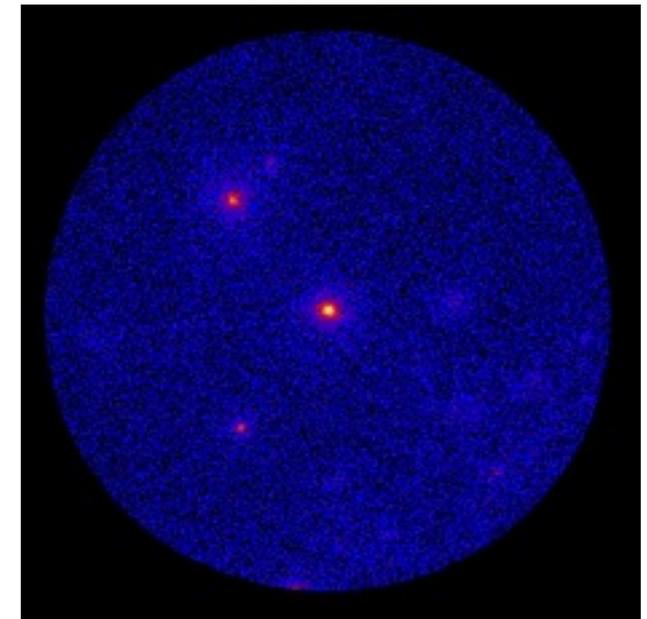
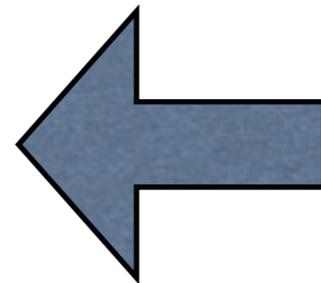
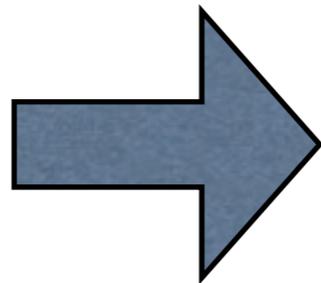
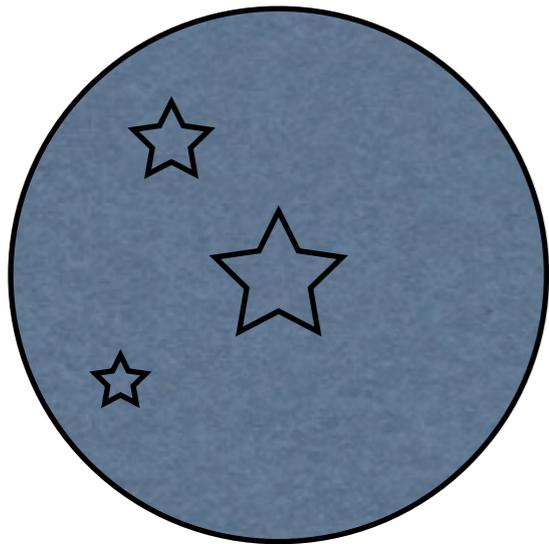
Our Data

Model 2



# A Graphical Example

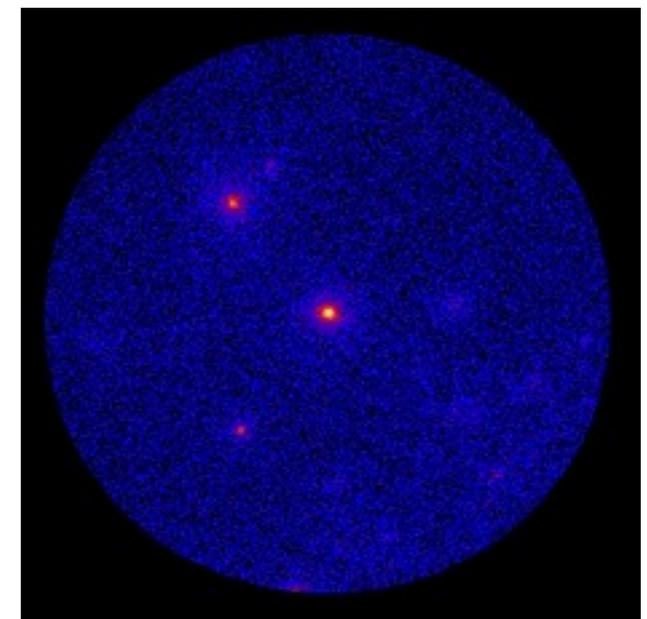
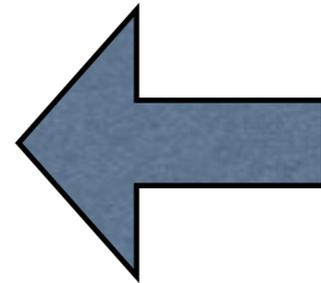
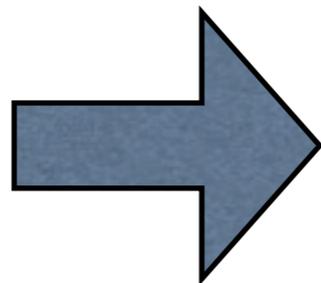
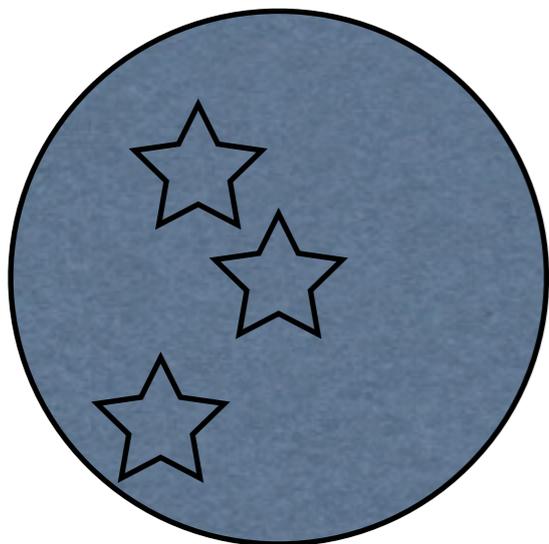
Model 3



Model 3 > Model 2

Our Data

Model 2



Like

# Likelihood Analysis: basics (I)

- The likelihood  $\mathcal{L}$  is the probability of obtaining your data given an input model.
- In our case, the input model is the distribution of gamma-ray sources on the sky and includes their intensity and spectra.
- One will maximize  $\mathcal{L}$  to get the best match of the model to the data. Given a set of data, one can bin them in multidimensional (energy, sky pixels, ...) bins.
- $\mathcal{L}$  is the product of the probabilities of observing the detected counts in each bin.

$$\mathcal{L} = \prod p_k$$

Like

# Likelihood Analysis: basics (2)

- The number of counts in each bin is small and thus is characterized by the Poisson distribution.
- $\mathcal{L}$  is the product of the probabilities of observing  $n_k$  counts in each bin when the number of counts predicted by the model is  $m_k$

$$\mathcal{L} = \prod_k \frac{m_k^{n_k} e^{-m_k}}{n_k!}$$

- $\mathcal{L}$  can be rewritten as

$$\mathcal{L} = \prod_k e^{-m_k} \prod_k \frac{m_k^{n_k}}{n_k!} = e^{-N_{pred}} \prod_k \frac{m_k^{n_k}}{n_k!}$$

Like

# Likelihood Analysis: basics (3)

- If we let the bin sizes get infinitesimally small, then  $n_k = (0 \text{ or } 1)$  and we are left with a product running over the number of photons:

$$\mathcal{L} = e^{-N_{pred}} \prod_i m_i$$

- (This is the *unbinned* likelihood method).
- It's easier to handle the logarithm of  $\mathcal{L}$  so we usually maximize:

$$\log \mathcal{L} = \sum_i \log(m_i) - N_{pred}$$

# Maximize This!

$$\log \mathcal{L} = \sum_i \log(m_i) - N_{pred}$$

$$\log \mathcal{L} = \sum_k n_k \log(m_k) - \sum_k m_k$$

The first term increases as the model predicts counts in bins where they actually occur and the second term demands that model counts be parsimoniously allocated.

Now, we need to quantify the significance.

Like

# Likelihood Analysis: basics (4)

The Test Statistic is defined as:

$$TS = -2 \log \left( \frac{\mathcal{L}_{\max,0}}{\mathcal{L}_{\max,1}} \right)$$

- Where  $\mathcal{L}_{\max,0}$  is the maximum likelihood value for a model without an additional source (the 'null hypothesis') and  $\mathcal{L}_{\max,1}$  is the maximum likelihood value for a model with the additional source at a specified location.
- In the limit of a large number of counts, Wilk's Theorem states that the TS for the null hypothesis is asymptotically distributed as  $X_n^2$  where  $n$  is the number of parameters characterizing the additional source.
- As a basic rule of thumb, the square root of the TS is approximately equal to the detection significance for a given source.

Like

# Likelihood Analysis: basics (5)

- The source model considered is

$$S(E, \hat{p}, t) = \sum_i s_i(E, t) \delta(\hat{p} - \hat{p}_i) + S_G(E, \hat{p}) + S_{eg}(E, \hat{p}) + \sum_l S_l(E, \hat{p}, t),$$

Point Sources      Galactic & EG Diffuse Sources      Other Sources

- This model is folded with the Instrument Response Functions (IRFs) to obtain the predicted number of counts in the measured quantity space ( $E', p', t'$ ):

$$M(E', \hat{p}', t) = \int_{\text{SR}} dE d\hat{p} R(E', \hat{p}', t; E, \hat{p}) S(E, \hat{p}, t)$$

Like

# Likelihood Analysis: basics (6)

- The IRF is defined as

$$R(E', \hat{p}', E, \hat{p}, \vec{L}(t)) = \frac{d\sigma}{dE' d\hat{p}'}(E, \hat{p}, \vec{L}(t))$$

- where  $E$  is the true energy of the photon,  $E'$  is the measured energy,  $\hat{p}_{\text{hat}}$  is the true direction and  $\hat{p}'_{\text{hat}}$  is the measured direction.  $L(t)$  contains all of the information about the telescope location and state.

Like

# Likelihood Analysis: basics (7)

It is often assumed that the IRF may be decomposed into three functions or matrices:

$$R(E', \hat{p}', E, \hat{p}, \vec{L}(t)) = D(E'; E, \hat{p}, \vec{L}(t))P(\hat{p}'; E, \hat{p}, \vec{L}(t))A(E, \hat{p}, \vec{L}(t)) \quad (2)$$

where

- $A(E, \hat{p}, \vec{L}(t))$  is the effective area, that is the projected area of the detector multiplied by its efficiency, that means the probability that a photon will react in the detector and produce a recognizable shower of particles.
- $D(E'; E, \hat{p}, \vec{L}(t))$  is the energy dispersion which reflects the imperfection of the energy measurement.
- $P(\hat{p}'; E, \hat{p}, \vec{L}(t))$  is the point spread function which reflects the imperfection of the measured direction.

Like

# Likelihood Analysis: basics (8)

- So, we perform the integral

$$M(E', \hat{p}', t) = \int_{\text{SR}} dE d\hat{p} R(E', \hat{p}', t; E, \hat{p}) S(E, \hat{p}, t)$$

- over the Source Region, i.e. the sky region encompassing all source contributions to the Region-of-interest (ROI). In the standard analysis, only steady sources are assumed:

$$S(E, \hat{p}, t) \rightarrow S(E, \hat{p})$$

# Likelihood Analysis: basics (9)

- The function to maximize is:

$$\log \mathcal{L} = \sum_j \log M(E'_j, \hat{p}'_j, t_j) - N_{\text{pred}}$$

- where the sum is performed over photons in the ROI.
- The predicted number of counts is:

$$N_{\text{pred}} = \int_{\text{ROI}} dE' d\hat{p}' dt M(E', \hat{p}', t)$$

- To save CPU time, a model-independent quantity called the exposure map (or cube) is precomputed:

$$\varepsilon(E, \hat{p}) \equiv \int_{\text{ROI}} dE' d\hat{p}' dt R(E', \hat{p}', t; E, \hat{p})$$

- Then:

$$N_{\text{pred}} = \int_{\text{SR}} dE d\hat{p} S(E, \hat{p}) \varepsilon(E, \hat{p})$$

# The One Slide Summary

- Observed a photon from a location, at a time, with an energy.

- Assume a model:

$$S(E, \hat{p}, t) = \sum_i s_i(E, t) \delta(\hat{p} - \hat{p}_i) + S_G(E, \hat{p}) + S_{\text{eg}}(E, \hat{p}) + \sum_l S_l(E, \hat{p}, t),$$

Point Sources      Galactic & EG Diffuse Sources      Other Sources

- Calculated the probability of that photon being detected assuming our model:

$$M(E', \hat{p}', t) = \int_{\text{SR}} dE d\hat{p} R(E', \hat{p}', t; E, \hat{p}) S(E, \hat{p}, t)$$

- Calculate the total number of predicted counts assuming our model.

$$N_{\text{pred}} = \int_{\text{SR}} dE d\hat{p} S(E, \hat{p}) \varepsilon(E, \hat{p})$$

- Adjust the model until this is maximized:

$$\log \mathcal{L} = \sum_j \log M(E'_j, \hat{p}'_j, t_j) - N_{\text{pred}}$$

- Calculate the TS:

$$TS = -2 \log \left( \frac{\mathcal{L}_{\text{max},0}}{\mathcal{L}_{\text{max},1}} \right)$$

# Overview of Tutorials

- Over the next few days, we'll show you how to do this practically with LAT data.
- **Today:** go over the basics of downloading, exploring and selecting data.
- **Goal:** basic understanding of what LAT data is and how the analysis is structured. Make sure you can run the tools.
- **Wednesday:** calculate exposures and live time. Perform a simple likelihood analysis.
- **Goal:** Successfully run a likelihood analysis
- **Thursday:** work with a more complicated region.
- **Goal:** Successfully find a convergent, significant model

# Overview of Tutorials

- Over the next few days, we'll show you how to practically do this with LAT data.
- **Friday:** Compare different models. Extract information from the likelihood fit.
- **Goal:** able to determine the best spectral shape for sources. Can extract needed information from the models.
- **Saturday:** Advanced analysis techniques: simulations, model testing, lightcurves, upper limits, binning, SEDs.
- **Goal:** Can make the tools do what you want.

# Homework

- Don't fall behind because you will be working on your own during the night.
- Some steps of the likelihood take a long time so you'll need to run these while at dinner, sleeping etc.
- Today: Pick an analysis (source, region, measurement) that you are interested in working on this week. Tonight, you'll start working on the analysis.